

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

## (i) APPLICANT:

- (A) NAME: LONZA AG
- (B) STREET: Muenchensteinerstrasse 38
- (C) CITY: Basle
- (E) COUNTRY: Switzerland
- (F) POSTAL CODE: 4002

(ii) TITLE OF INVENTION: Process for the preparation of (S)- or (R)-3,3,3-trifluoro-2-hydroxy-2-methylpropionic acid

(iii) NUMBER OF SEQUENCES: 14

## (iv) COMPUTER-READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (c) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

## (2) INFORMATION FOR SEQ ID NO: 1:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1442 base pairs
- (B) TYPE: Nucleotide
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: Genomic DNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

## (vi) ORIGIN:

- (A) ORGANISM: Klebsiella oxytoca
- (B) STRAIN: PRS1
- (C) INDIVIDUAL/ISOLATE: PRS1

## (vii) PROVENANCE:

- (B) CLONE(S): pPRS2a

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: join(197..1181)
- (D) OTHER INFORMATION: /product= "amidase"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

CCCGGGA	ACT	CCATGTGGCC	GTGATCCTGG	TCGAGCAGGA	TATTGCGATG	ATCCAGCGGG	60
CCGCACAGCG	CTGTGCGGTA	ATG3ATAAAG	GCCTGGTTGT	AGAAAACGCTG	ACCCAACAAC		120
AGCTCTCTGA	TGATCTTTTA	ATGCGTCGTC	ATCTGGCTCT	GTAAGTAAAC	GCTATAAAATT		180
ACGTGGAGAA	TAACAT	ATG	AAA	TGG	TIG	GAA	229
	Met	Lys	Trp	Leu	Glu	Glu	
	1			5		10	

[illegible]

CGC Arg	GGT Gly	GTT Val	GGT Gly 15	GCC Ala	GGG Gly	CGT Arg	AAA Lys	CCG Pro 20	GTA Val	ACG Thr	CAT His	CAC His	CTG Leu 25	ACG Thr	GAA Glu	277
GAA Glu	ATG Met	CAA Gln 30	AAA Lys	GAG Glu	TTT Phe	CAT His	TAC Tyr 35	ACC Thr	ATT Ile	GGC Gly	CCT Pro	TAT Tyr 40	TCC Ser	ACA Thr	CCC Pro	325
GTC Val	CTG Leu 45	ACC Thr	ATC Ile	GAA Glu	CCC Pro	GGT Gly 50	GAC Asp	CGG Arg	ATT Ile	ATT Ile	GTC Val 55	GAC Asp	ACT Thr	CGA Arg	GAT Asp	373
GCT Ala 60	TTT Phe	GAA Glu	GGT Gly	GCT Ala	ATC Ile 65	AAT Asn	TCG Ser	GAA Glu	CAG Gln	GAT Asp 70	ATT Ile	CCG Pro	AGC Ser	CAG Gln	TTG Leu 75	421
CTA Leu	AAA Lys	ATG Met	CCC Pro	TTT Phe 80	CTC Leu	AAC Asn	CCA Pro	CAA Gln	AAC Asn 85	GGA Gly	CCG Pro	ATC Ile	ATG Met	GTC Val 90	AAT Asn	469
GGC Gly	CGC Ala	GAG Glu	AAA Lys 95	GGT Gly	GAT Asp	GTG Val	CTC Leu 100	GCT Ala	GTG Val	TAT Tyr	ATC Ile	GAA Glu	TCC Ser 105	ATG Met	TTG Leu	517
CCC Pro	CGC Arg	GGC Gly 110	GTT Val	GAT Asp	CCC Pro	TAC Tyr	GGC Gly 115	ATC Ile	TGC Cys	GCC Ala	ATG Met	ATT Ile 120	CCG Pro	CAT His	TTT Phe	565
GGC Gly 125	GGA Gly	CTG Leu	ACC Thr	GGG Gly	ACC Thr	GAC Asp 130	CTG Leu	ACG Thr	GCC Ala	ATG Met	CTC Leu 135	AAT Asn	GAT Asp	CCG Pro	CTG Leu	613
CCA Pro 140	GAA Glu	AAG Lys	GTG Val	CGC Arg	ATG Met 145	ATT Ile	AAA Lys	CTC Leu	GAC Asp	AGT Ser 150	GAA Glu	AAG Lys	GTC Val	TAC Tyr	TGG Trp 155	661
AGC Ser	AAA Lys	CGC Arg	CAT His	ACG Thr 160	CTT Leu	CCC Pro	TAT Tyr	AAA Lys	CCC Pro 165	CAT His	ATT Ile	GGC Gly	ACC Thr	TTG Leu 170	AGC Ser	709
GTA Val	TCG Ser	CCA Pro	GAA Glu 175	ATT Ile	GAC Asp	TCA Ser	ATC Ile	AAT Asn 180	TCA Ser	CTG Leu	ACG Thr	CCA Pro	GAC Asp 185	AAT Asn	CAC His	757
GGC Gly	GGG Gly	AAT Asn 190	ATG Met	GAT Asp	GTG Val	CCG Pro	GAT Asp 195	ATA Ile	GGA Gly	CCA Pro	GGG Gly	AGT Ser 200	ATT Ile	ACC Thr	TAT Tyr	805
CTG Leu 205	CCG Pro	GTA Val	CGT Arg	GCG Ala	CCT Pro	GGA Gly 210	GGC Gly	CGC Arg	CTG Leu	TTT Phe 215	ATT Ile	GGT Gly	GAT Asp	GCC Ala	CAT His	853
GCT Ala 220	TGT Cys	CAG Gln	GGT Gly	GAT Asp	GGT Gly 225	GAG Glu	ATT Ile	TGC Cys	GGG Gly	ACC Thr 230	GCA Ala	GTA Val	GAG Glu	TTT Phe	GCC Ala 235	901
TCA Ser	ATC Ile	ACC Thr	ACC Thr	ATC Ile 240	AAA Lys	GTC Val	GAT Asp	TTG Leu	ATC Ile 245	AAG Lys	AAC Asn	TGG Trp	CAG Gln	CTT Leu 250	TCC Ser	949
TGG Trp	CCA Pro	CGA Arg	ATG Met 255	GAG Glu	AAT Asn	GCC Ala	GAA Glu	AAT Asn 260	ATT Ile	ATG Met	AGT Ser	ATT Ile	GGC Gly 265	AGT Ser	GCA Ala	997

CGT CCG CTG GAG GAT GCG ACG CGA ATT GCA TAT CGC GAC TTA ATT TAC Arg Pro Leu Glu Asp Ala Thr Arg Ile Ala Tyr Arg Asp Leu Ile Tyr 270 275 280	1045
TGG CTG GTA GAA GAC TTT GGC TTC GAA CAA TGG GAT GCC TAC ATG CTT Trp Leu Val Glu Asp Phe Gly Phe Glu Gln Trp Asp Ala Tyr Met Leu 285 290 295	1093
CTG AGT CAA TGC GGC AAA GTG CGG CTG GGC AAC ATG GTC GAC CCC AAA Leu Ser Gln Cys Gly Lys Val Arg Leu Gly Asn Met Val Asp Pro Lys 300 305 310 315	1141
TAC ACC GTT GGC GCG ATG CTG AAC AAA AAC CTG TTA GTT TAGTAGGAAT Tyr Thr Val Gly Ala Met Leu Asn Lys Asn Leu Leu Val 320 325	1190
AACTAACCGG TGAACATTAC CCGGATGTAG ATCGGGGTAA TGTGTAAGTT CAAACAATCG	1250
CTATTTTAA CAGCTAAAGC AGGTGCATAT GGGGCCAGAT ACACCCATCA ATATTGGTTT	1310
ACTTTACTCC TTCAGCGGAG TGACGGCGGC ACAGAGTTG TCACAATGGC GCGGAGCAAC	1370
CCAGGCTATT GCCGAAATTA ATCAAAATGG CGGCATCAAC GGCAGACCAC TCAATGCAAT	1430
TCATTTGGAT CC	1442

## (2) INFORMATION FOR SEQ ID NO: 2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 328 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Lys Trp Leu Glu Glu Ser Ile Met Ala Lys Arg Gly Val Gly Ala 1 5 10 15
Gly Arg Lys Pro Val Thr His His Leu Thr Glu Glu Met Gln Lys Glu 20 25 30
Phe His Tyr Thr Ile Gly Pro Tyr Ser Thr Pro Val Leu Thr Ile Glu 35 40 45
Pro Gly Asp Arg Ile Ile Val Asp Thr Arg Asp Ala Phe Glu Gly Ala 50 55 60
Ile Asn Ser Glu Gln Asp Ile Pro Ser Gln Leu Leu Lys Met Pro Phe 65 70 75 80
Leu Asn Pro Gln Asn Gly Pro Ile Met Val Asn Gly Ala Glu Lys Gly 85 90 95
Asp Val Leu Ala Val Tyr Ile Glu Ser Met Leu Pro Arg Gly Val Asp 100 105 110
Pro Tyr Gly Ile Cys Ala Met Ile Pro His Phe Gly Gly Leu Thr Gly 115 120 125
Thr Asp Leu Thr Ala Met Leu Asn Asp Pro Leu Pro Glu Lys Val Arg 130 135 140

Met Ile Lys Leu Asp Ser Glu Lys Val Tyr Trp Ser Lys Arg His Thr  
145 150 155 160

Leu Pro Tyr Lys Pro His Ile Gly Thr Leu Ser Val Ser Pro Glu Ile  
165 170 175

Asp Ser Ile Asn Ser Leu Thr Pro Asp Asn His Gly Gly Asn Met Asp  
180 185 190

Val Pro Asp Ile Gly Pro Gly Ser Ile Thr Tyr Leu Pro Val Arg Ala  
195 200 205

Pro Gly Gly Arg Leu Phe Ile Gly Asp Ala His Ala Cys Gln Gly Asp  
210 215 220

Gly Glu Ile Cys Gly Thr Ala Val Glu Phe Ala Ser Ile Thr Thr Ile  
225 230 235 240

Lys Val Asp Leu Ile Lys Asn Trp Gln Leu Ser Trp Pro Arg Met Glu  
245 250 255

Asn Ala Glu Asn Ile Met Ser Ile Gly Ser Ala Arg Pro Leu Glu Asp  
260 265 270

Ala Thr Arg Ile Ala Tyr Arg Asp Leu Ile Tyr Trp Leu Val Glu Asp  
275 280 285

Phe Gly Phe Glu Gln Trp Asp Ala Tyr Met Leu Leu Ser Gln Cys Gly  
290 295 300

Lys Val Arg Leu Gly Asn Met Val Asp Pro Lys Tyr Thr Val Gly Ala  
305 310 315 320

Met Leu Asn Lys Asn Leu Leu Val  
325

## (2) INFORMATION FOR SEQ ID NO: 3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not known
- (D) TOPOLOGY: not known

## (ii) MOLECULE TYPE: peptide

## (vi) ORIGIN:

- (B) STRAIN: PRS1
- (C) INDIVIDUAL/ISOLATE: PRS1

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Met Lys Trp Leu Glu Glu Ser Ile Met Ala Lys Arg Gly Val Gly Ala  
1 5 10 15

Ser Arg Lys Pro  
20

## (2) INFORMATION FOR SEQ ID NO: 4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not known
- (D) TOPOLOGY: not known

(ii) MOLECULE TYPE: peptide

(vi) ORIGIN:

(B) STRAIN: PRS1

(C) INDIVIDUAL/ISOLATE: PRS1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Val Tyr Trp Ser Lys  
1 5

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 13 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: not known

(D) TOPOLOGY: not known

(ii) MOLECULE TYPE: peptide

(vi) ORIGIN:

(B) STRAIN: PRS1

(C) INDIVIDUAL/ISOLATE: PRS1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

Lys Pro Val Thr His His Leu Thr Glu Glu Met Gln Lys  
1 5 10

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: not known

(D) TOPOLOGY: not known

(ii) MOLECULE TYPE: peptide

(vi) ORIGIN:

(B) STRAIN: PRS1

(C) INDIVIDUAL/ISOLATE: PRS1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Tyr Thr Val Gly Ala Met Leu Asn Lys  
1 5

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 14 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: not known

(D) TOPOLOGY: not known

(ii) MOLECULE TYPE: peptide

(vi) ORIGIN:

(B) STRAIN: PRS1

(C) INDIVIDUAL/ISOLATE: PRS1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

Met Glu Asn Ala Glu Asn Ile Met Ser Ile Gly Ser Ala Arg  
1 5 10

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: not known

(D) TOPOLOGY: not known

(ii) MOLECULE TYPE: peptide

(vi) ORIGIN:

(B) STRAIN: PRS1

(C) INDIVIDUAL/ISOLATE: PRS1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Trp Leu Glu Glu Ser Ile Met Ala Lys  
1 5

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: not known

(D) TOPOLOGY: not known

(ii) MOLECULE TYPE: peptide

(vi) ORIGIN:

(B) STRAIN: PRS1

(C) INDIVIDUAL/ISOLATE: PRS1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

Met Pro Phe Leu Asn Pro Gln Asn Gly Pro Ile Met Val Asn Gly Ala  
1 5 10 15

Glu Lys

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 amino acids

(B) TYPE: amino acid

- (C) STRANDEDNESS: not known  
(D) TOPOLOGY: not known

(ii) MOLECULE TYPE: peptide

(vi) ORIGIN:

- (B) STRAIN: PRS1  
(C) INDIVIDUAL/ISOLATE: PRS1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

Asp Ala Phe Glu Gly Ala Ile Asn Ser Glu Gln Asp Ile Pro Ser Gln  
1 5 10 15

Leu Leu Lys

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: not known  
(D) TOPOLOGY: not known

(ii) MOLECULE TYPE: peptide

(vi) ORIGIN:

- (B) STRAIN: PRS1  
(C) INDIVIDUAL/ISOLATE: PRS1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

Glu Phe His Tyr Thr Ile Gly Pro Tyr Ser Thr Pro Val Leu Thr Ile  
1 5 10 15

Glu Pro Gly Asp Arg  
20

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: not known  
(D) TOPOLOGY: not known

(ii) MOLECULE TYPE: peptide

(vi) ORIGIN:

- (B) STRAIN: PRS1  
(C) INDIVIDUAL/ISOLATE: PRS1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

Leu Phe Ile Gly Asp Ala His Ala Glu Gln Gly Asp Gly Glu Ile Glu

1 5 10 15

Gly Thr Ala Val Glu Phe Ala  
20



(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 14 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: not known  
(D) TOPOLOGY: not known

## (ii) MOLECULE TYPE: peptide

(vi) ORIGIN:

- (B) STRAIN: PRS1  
(C) INDIVIDUAL/ISOLATE: PRS1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

Gly Asp Val Leu Ala Val Tyr Ile Glu Ser Met Leu Pro Arg  
1 5 10

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: not known  
(D) TOPOLOGY: not known

## (ii) MOLECULE TYPE: peptide

(vi) ORIGIN:

- (C) INDIVIDUAL/ISOLATE: PRS1

(vii) PROVENANCE:

- (B) CLONE(S) : PRS1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

Gly Val Asp Pro Tyr Gly Ile Glu Ala Met Ile Pro His Phe Gly Gly  
1 5 10 15

Leu Thr Gly Thr Asp Leu Thr Ala Met Leu Asn Asp Gin Leu Gin Pro  
20 25 30

**Lys**